人脸识别系统

一、 课题介绍

随着计算机以及相关技术的发展，人们对智能化设备的兴趣逐渐提高。人们希望能够仅仅使用人脸，就能够表明身份，获得相关的权限。人脸识别被期望用于安防领域，从小区楼栋的防盗门到公共环境中摄像头图像的人脸识别，都需要人脸识别技术的支持。有时，对于图像中大量的人物特征不清晰，难以使用人眼进行辨别，可以使用人脸识别技术对图像中的人脸进行识别。人脸识别技术对于图像中大量人脸的识别具有优势。

二、 相关工作

人脸识别系统主要是由人脸检测和人脸识别两个部分组成。

人脸检测的过程是指向系统中输入一张含有人脸的图像，系统自动检测出图像中的人脸，并且能够正确地标识出人脸区域。图像技术经过长期地发展，出现了许多人脸检测的算法。人脸检测算法的效率和正确率都在不断地提升。二十一世纪初，人脸检测算法还是基于肤色的检测。然而最近几年，基于 HAAR 特征的人脸检测算法得到了广泛地应用，由于它较高的效率和正确率。经过 HAAR 特征地提取，再把提取的结果输入到分类器中就能够实现人脸检测过程。

人脸识别特指利用分析比较人脸视觉特征信息进行身份鉴别的计算机技术。人脸识别是一项热门的计算机技术研究领域；它属于生物特征识别技术，是对生物体（一般特指人）本身的生物特征来区分生物体个体。进行人脸图像识别研究具有很大的使用价值。如同人的指纹一样，人脸也具有唯一性，也可用来鉴别一个人的身份。现在已有实用的计算机自动指纹识别系统面世，并在安检等部门得到应用，但还没有通用成熟的人脸自动识别系统出现。人脸图像的自动识别系统较之指纹识别系统、DNA 鉴定等更具方便性，因为它取样方便，可以不接触目标就进行识别，从而开发研究的实际意义更大。并且与指纹图像不同的是，人脸图像受很多因素的干扰:人脸表情的多样性;以及外在的成像过程中的光照，图像尺寸，旋转，姿势变化等。使得同一个人，在不同的环境下拍摄所得到的人脸图像不同，有时更会有很大的差别，给识别带来很大难度。因此在各种干扰条件下实现人脸图像的识别，也就更具有挑战性。当前大多数人脸识别算法是基于统计的方法，也就是需要大量的样本如特征脸方法（PCA）、Fisher 脸方法、奇异值分解方法、神经网络方法和支持向量机等，在这些方法中影响最大的是 PCA 方法和神经网络方法。我们主要用的是卷积神经网络的方法来进行人脸识别。卷积神经网络是一个受生物视觉启发、以最简化预处理操作为目的的多层感知器的变形，

本质是一个前向反馈神经网络，卷积神经网络与多层感知器的最大区别是网络前几层由卷积层和池化层交替级联组成，模拟视觉皮层中用于高层次特征提取的简单细胞和复杂细胞交替级联结构。卷积层的神经元对前一层输入的一部分区域（称为局部感受野，区域之间有重叠）有响应，提取输入的更高层次特征；池化层的神经元对前一层输入的一部分区域（区域之间无重叠）求平均值或最大值，抵抗输入的轻微形变或者位移。卷积神经网络的后几层一般是若干个全连接层和一个分类器构成的输出层。卷积神经网络用于人脸识别是一种基于特征的方法，区别于传统的人工特征提取和针对特征的高性能分类器设计，它的优点是通过逐层卷积降维进行特征提取，然后经过多层非线性映射，使网络可以从未经特殊处理的训练样本中，自动学习形成适应该识别任务的特征提取器和分类器，该方法降低了对训练样本的要求，而且网络的层数越多，学习到的特征更具有全局性。

三、 方法实现

人脸检测部分：

基于 HAAR 特征的人脸检测是对整张图像进行计算，找到可能是人脸的区域。将检测结果输入到分类器中，分类出人脸图像。再根据分类出的人脸图像，找到原图像中的人脸区域。

1. 输入图像

向系统中输入包含有人脸的图像，将图像中转化成矩阵。获得图像矩阵的大小等信息。

2. 图像灰度化以及图像增强

将原图像进行灰度化，灰度化可以保留图像的大部分特征（除颜色特征外），并且能够减少大量的存储空间，使得在图像存储和计算中能够有较高的效率。常用的灰度化方法是将图像颜色三通道求均值形成单通道图像。观察灰度化的结果，图像中的大部分特征都能够被很好地保留。将灰度化后的图像直接用于检测，检测效果并不是很好。于是，考虑使用图像增强的方法对图像再次进行预处理，增强图像特征与背景的对比度，提高检测的效果。在图像增强时，使用直方图均衡化的方法。直方图均衡化是将直方图中各个区域的间隔增大，使得每一个主要特征的区别度都增大，增加不同主要特征的差异。

3. 定义分类器并使用 HAAR 特征检测

HAAR 特征主要有原始矩形特征、边缘特征和线特征。利用 HAAR 特征在原图像上进行滑动检测，将白色区域内的像素值之和减去黑色区域内的像素值之和，得到的数值作为人脸特征化的一个数值。通过这样的方法对原图像进行处理之后，人脸区域计算出的数值与背景区域计算出的数值存在一定的差异，而人脸区域的数值接近。接下来可以使用分类器对整幅图像计算出的数值进行分类，以获得人脸区域。AdaBoost 算法是由一系列弱分类器经过级联形成的一个强分类器，分类效果比较理想。AdaBoost 算法的结构类似于二叉决策树，每一个节点都是一个分类器。训练时，将所有带有标签的数据输入算法中，经过每一个节点的计算和分类，不断调整每一个节点分类器的权值，最终使得分类器对特征的分类正确率达到最高。经过大量实验数据的训练，可以得到一个性能不错的分类器。对 HAAR 特征的特征值进行分类，就能够得到准确的人脸与非人脸分类结果。通过调节检测窗口可以得到不同的结果。

4. 粗略判断是否为人脸

经过 HAAR 特征的检测以及 AdaBoost 分类器的分类，结果中仍然有错误的分类结果。经过对这些分类结果的观察，发现人脸区域大小相近，而非人脸区域大小与人脸区域大小差距较大。于是，利用检测结果中图像的大小作为一个粗略的标准对结果再一次进行分类。经过这样的判断之后，错检的数量得到明显地减少，且运行时间并没有明显地增加。

5. 原图像中标识人脸区域

人脸检测的结果直接在原图像中标识出来，具有更好地表现效果。找到检测的人脸区域在原图像中的位置和检测图像的大小，根据图像大小在原图像中的位置修改像素值，形成一个明显的矩形框，将人脸标识出来。

6. 保存人脸图像

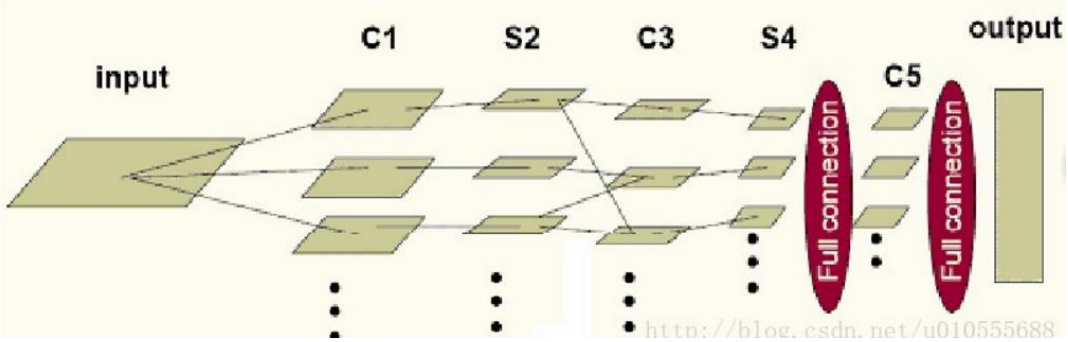
由于后续识别部分需要人脸图像，因此将人脸图像从原图中截取出来。再将图像矩阵转化成图像存储起来。

人脸识别部分：

人脸识别部分，目的是将单个人脸图片的身份信息识别出来，因此我们的主要的模块是卷积神经网络的构建、训练、测试以及优化等。

1.神经网络的构建

我们的神经网络主要用的是 cnn 卷积神经网络，它包括一个输入层，两个卷积层和两个下采样层、一个全连接层、一个输出层，基本结构如图：



1.1　输入层

我们的输入层的功能是直接将一个只包含单人脸图像输入到卷积神经网络中。经过对 28*28，32*32，36*36，40*40 大小的图片输入试验发现 36*36 的图片准确率最高。因此我们将输入图片的大小设为 36*36.

1.2　卷积层和下采样层

卷积层通过局部连接和权值共享的方法，模拟具有局部感受野的简单细胞，提取一些初级视觉特征的过程。局部连接指卷积层上的每一个神经元与前一层特征图中固定区域的神经元建立连接；权值共享指同一特征图中的神经元用一组相同的连接强度与前一层局部连接，可以减少网络训练参数，上述一组相同的连接强度即为一个特征提取器，在运算的过程中变现为一个卷积核，卷积核数值先随机初始化，最后由网络训练确定。

池化层模拟复杂细胞是将初级的视觉特征筛选并结合成更高级、抽象的视觉特征的过程，在网络中通过采样实现，经过池化层的采样后，输出特征图的数量不变，但是特征图的尺寸会变小，有减小计算复杂度、抵抗微小位移变化的作用。

两个卷积层和两个下采样层（也叫池化层）是交叉顺序排列的，即每个卷积层后面都要跟一个下采样层。在这一部分每个卷积层的卷积核大小都是 5*5，下采样层大小都是 2*2。第一个卷积层的卷积核数量是 10，第二个卷积层卷积核数量是 20。

1.3　全连接层

为了增强网络的非线性能力，同时限制网络规模的大小，网络在 4 个特征提

取层提取特征后，接入一个全连接层，该层的每一个神经元与前一层的所有神经元互相连接，同层神经元之间不连接。

1.4　输出层

输出层将神经网络的分类结果输出，输出层输出的是一个列向量，找出数字最大的一个结果，就是正确可能性最大的分类。

2.样本的采集及处理

卷积神经网络最重要的工作之一是对训练集和测试集采集及处理，只有对样本适当的采集和处理才能提高神经网络的准确率。

2.1　样本的采集

由于我们的实验目的是将一张照片中的人脸检测到并识别出来，这最直接的测试是全班同学合照中的人脸检测。因此，我们收集了 26 位同学再加上网上下载的 15 个人的最少十一张照片作为样本集，其中十张为训练集，一张为测试集。然后标准化 2 个集合中的人脸。

我们用的是“min-max”标准化灰度值，即对图片中每个像素点的灰度值标准化为[0,1]。用 x 和 x’分别表示当前和标准化后的灰度值，min 和 max 分别表示图片的最小和最大的灰度值，标准化公式为：

x’=（x-min）/(max-min)

2.2　样本的处理

在对采集到的样本直接训练后，测试集的正确率为 82.54%，但是在将几个人的合照放入代码实际测试时正确率却几乎为 0%。

由于最终测试结果并不理想，我们对训练集做了预处理，我们用了噪声扰动、对比度变换等方式对样本进行了数据增强。噪声包括高斯噪声，椒盐噪声，这些干扰的强度都是随机产生的，也就是说，每张图片每次训练的时候都不完全一样。这就增大了训练样本的容量，一定程度上解决了样本数量少的问题。

3.训练及测试

我们采用的是在实践中收敛速度较快的批量随机梯度下降法。迭代步长为 1，批处理块为 10，迭代次数为 1000，每次迭代会遍历训练集的所有批处理块，遍历完一个批处理块更新一次网络参数．更新公式为

$$w_{i+1}=\varepsilon \cdot w_i-\eta \cdot \left(\frac{\partial L}{\partial w_i}\right)_{D_i}. \tag{7}$$

其中：w_i 是当前参数，w_{i+1}是更新后的参数，ε 是动量，η 是学习速率，$\left(\frac{\partial L}{\partial w_i}\right)_{D_i}$ 是第 i 个批处理块 D_i 中误差对 w_i 偏导的平均值.

4.测试结果的处理

由于我们无法收集到所有同学的照片，因此为尽量减少误差，将训练集中没有的同学和检测出的错误人脸标记，我们引入一个阈值δ。由于卷积神经网络的输出是一个列向量，于是我们设 H 为这个列向量。

$$\sigma=\frac{\mathrm{P}}{\sum_{i=1}^{H} p_i}$$

P为向量H的最大一个值

$\sum_{i=1}^{H} p_i$ 为向量H各行之和

这个参数的主要功能就是判断训练集中是否有这个人的人脸图像，如果没有则标记“未识别”，以减小出错率。经过观察发现在 0.5~0.6 的时候正确率最高。

5.参数设置

神经网络的参数对神经网络的准确率具有最直接的影响，以下是神经网络参数介绍。

1)学习速率(learning_rate)：非常重要的一个参数，是运用 SGD 算法时梯度前面的系数，如果设得太大可能算法永远都无 法优化，太小会让算法优化太慢，并且有可能会掉入局部最优。

2）批次大小(batch_size)：将数据输入模型是按批次的，然 后计算这个批次所有样本的平均损失，也就是代价函数是所有样本的平均。因此批次的大小也会影响到对模型的优化程度 和速度。

3）训练步数(n_epochs)：遍历训练集的次数。

4）卷积层的卷积核个数(n_kerns)：卷积核的个数代表特征 的个数，卷积核的个数越多提取的特征就越多，可能最后分类会越准确。但是如果卷积核太多，

会增加参数的规模和计算复杂度。

5)池化大小(pool_size)：这里采用的是最大池化方法来选 取图像区域的最大值来作为该区域池化后的值，比如(2,2)为选取区域 2*2 的像素中的最大值。

由于采用的数据集并不大，本次训练将学习速率调成 1，批次大小设置成 10，训练步数为 1000，第一个卷积层的卷积核个数为 10，第二个为 20，池化大小为(2,2)。

四、 检测及识别效果

检测效果：

图 1 原图像

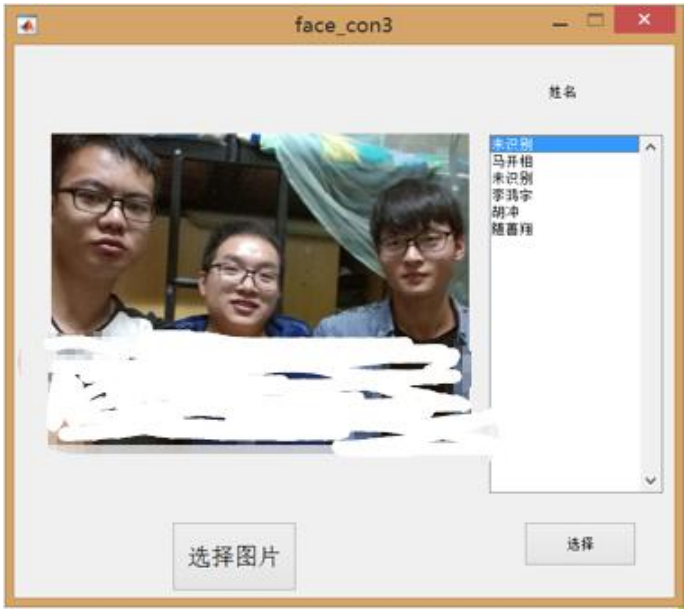
图 2 检测窗口每次增大 10%

图 3 检测窗口每次增大 5%

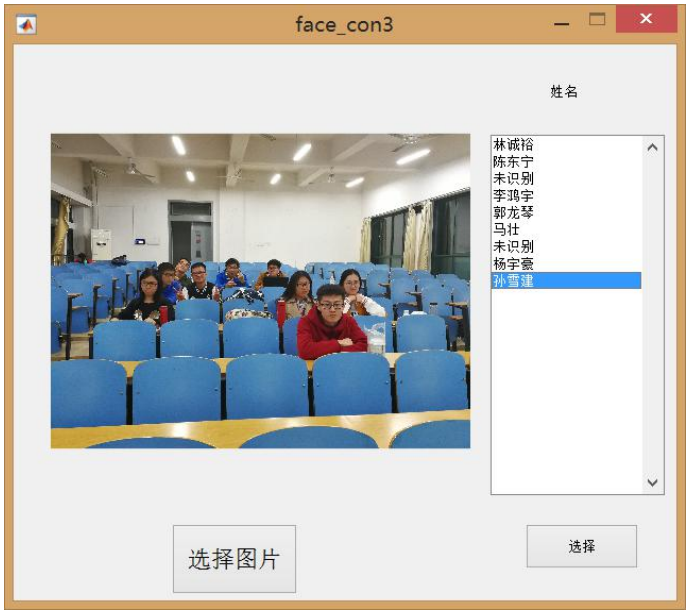
图 4　检测窗口每次增大 1%，结果未处理

图 5 检测窗口每次增大 1%，结果粗略处理

识别效果：



图中是运行结果，人脸检测识别出图中有 6 个人脸，多检测出三个错误人脸，这三个人脸识别结果有两个标记“未识别”有一个识别错误，检测对三个人脸，这三个人脸全部识别正确。



图中的人脸变得多了起来，总的识别正确率为 44.44%其中人脸检测出 8 个正确人脸一个错误人脸，在 8 个正确人脸中有 4 个识别正确，检测出的一个错误人脸成功标记了“未识别”。